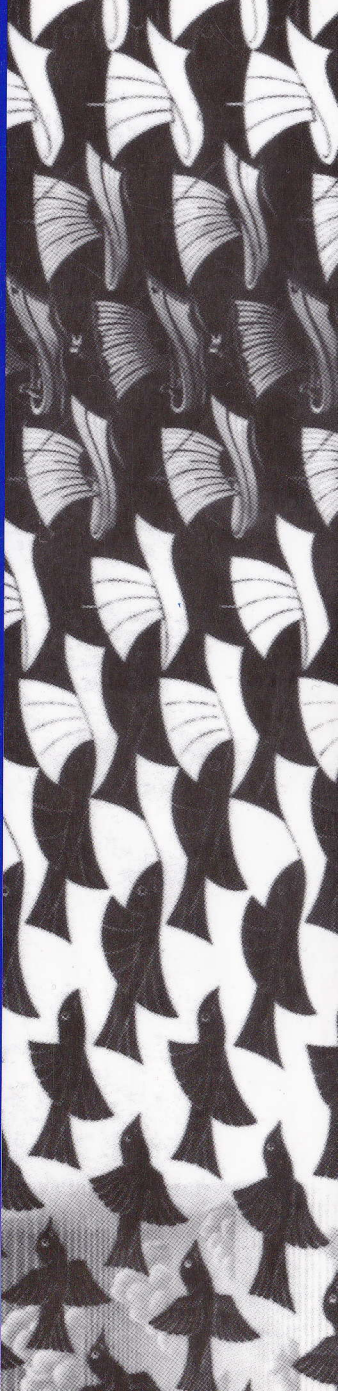
FILOSOFÍA

Razón y racionalidad

Jon Elster



Amorrortu/*editores*

Razón y racionalidad

Jon Elster

Amorrortu editores
Buenos Aires - Madrid

Biblioteca de filosofía
Reason and Rationality, Jon Elster
© Princeton University Press, 2008
Traducción: Horacio Pons

© Todos los derechos de la edición en castellano reservados por Amorrortu editores S.A., Paraguay 1225, 7° piso - C1057AAS Buenos Aires
Amorrortu editores España S.L., C/López de Hoyos 15, 3° izquierda - 28006 Madrid

www.amorrortueditores.com

La reproducción total o parcial de este libro en forma idéntica o modificada por cualquier medio mecánico, electrónico o informático, incluyendo fotocopia, grabación, digitalización o cualquier sistema de almacenamiento y recuperación de información, no autorizada por los editores, viola derechos reservados.

Queda hecho el depósito que previene la ley n° 11.723

Industria argentina. Made in Argentina

ISBN 978-950-518-360-9 (Argentina)
ISBN 978-0-691-13900-5, Princeton, NJ, edición original

Elster, Jon
Razón y racionalidad - 1ª ed. - Buenos Aires : Amorrortu, 2014.
72 p.; 20x12 cm. - (Colección Filosofía)

Traducción de: Horacio Pons

ISBN 978-950-518-360-9

1. Filosofía. 2. Ciencias Sociales. I. Pons, Horacio, trad. II. Título.
CDD 190

Impreso en los Talleres Gráficos Color Efe, Paso 192, Avellaneda, provincia de Buenos Aires, en marzo de 2014.

Tirada de esta edición: 2.000 ejemplares.

Índice

9 Razón y racionalidad

59 Referencias bibliográficas

67 Las lecciones inaugurales del Collège de France

En los enfoques analíticos de los comportamientos humanos, una misma palabra latina, «*ratio*», está en la raíz de dos tradiciones intelectuales muy diferentes pero a la vez ligadas entre sí.

Una de ellas es la que opone la razón a las pasiones y, en época más reciente, a los intereses. El tratado de Séneca *De la cólera*, por ejemplo, se organiza alrededor de la oposición entre razón y pasión, en tanto que los moralistas franceses del siglo XVII le suman el interés. En un texto célebre, La Bruyère resume así sus relaciones mutuas: «Nada cuesta menos a la pasión que ponerse por encima de la razón: su gran triunfo es imponerse al interés».[1] La idea de la razón está íntimamente ligada a la del bien común.

[1] La Bruyère (1990), pág. 150. [El lector encontrará las referencias bibliográficas al final del volumen, págs. 59 y sigs. (*N. del E.*)]

La otra es la noción, aún más reciente, de la elección racional, que se opone a las diversas formas de la irracionalidad. El actor racional es aquel que actúa por razones suficientes. Esas razones son las creencias o los deseos a la luz de los cuales la acción aparece como apropiada, en un sentido al que deberé volver *in extenso*. Con frecuencia, pero erróneamente, la idea de racionalidad se vincula a la del bien privado o el interés, en el sentido de los moralistas. Quien persigue el bien común puede —e incluso debe— hacerlo de manera racional.

Actuar conforme a la razón, en singular, y actuar por buenas razones, en plural, son dos cosas diferentes, por cuanto la razón es objetiva, mientras que las razones son subjetivas. Desde el punto de vista externo, una política se puede evaluar como acorde o no a la razón. Desde el punto de vista interno, una acción puede evaluarse como racional o no racional.[2] De esta diferencia se deduce que, a los fines explicativos, sólo se puede apelar a la racionalidad. Únicamente en la medida en que

[2] Sobre esta distinción, véase Williams (1981).

el agente hace suyas las exigencias de la razón pueden estas generar, y llegado el caso explicar, comportamientos específicos.

A pesar de ser diferentes, las dos normas tienen un obstáculo común, a saber: las pasiones.[3] Tie-

[3] Aunque clásica, esta oposición ha suscitado recientemente diversas objeciones. En lo concerniente a la relación entre las emociones y la racionalidad, véase en especial Damasio (2000), según el cual los individuos que sufren una lesión cerebral tienen, a la vez, una racionalidad deficiente y una menor capacidad de experimentar emociones. Ahora bien, no está demostrado, al parecer, que se trate de un lazo causal y no de una simple correlación. En cuanto a la relación entre las emociones y la razón, véase sobre todo Rudenfeld (2001), quien sostiene que la razón necesita la ayuda de la pasión para superar el interés. Tal fue, asimismo, la idea que Stanislas de Clermont-Tonnerre expuso en la Asamblea Constituyente: «La anarquía es una travesía espantosa pero necesaria, y el único momento en que puede llegarse a un nuevo orden de cosas. No será en tiempos de calma cuando se tomen medidas uniformes» (*AP*, 9, pág. 461). La noche del 4 de agosto pudo sugerir la misma conclusión. Los debates de la Asamblea Constituyente contienen, sin embargo, ambigüedades que complican gravemente las cosas; es preciso estar de acuerdo con Tocqueville, por ejemplo, cuando escribe que esa «noche (. . .) fue el producto combinado, pero en dosis

nen también un componente común, que es la idea de actuar sobre la base de creencias bien fundadas. Por último, comparten el hecho de que el actor las hace objeto de cierta deferencia. El origen y la naturaleza de esta última no son los mismos, pero en ambos casos se trata de deferencia para con una fuente de normatividad.[4] El funcionamiento de los mecanismos de deferencia es complejo. Baste con decir, por el momento, que su efecto consiste a veces en subvertir el objeto de la deferencia.

Se replicará tal vez que comparar un principio concerniente a la filosofía política normativa con otro que atañe a la explicación del comportamiento individual es una iniciativa endeble. Una respuesta, modesta pero suficiente, a esta objeción consistiría en decir que, dadas las confusiones habituales al respecto, el esclarecimiento conceptual es una tarea que por sí misma merece ser llevada adelante.

---

imposibles de precisar, del miedo y el entusiasmo» (2004, pág. 593).

[4] Tomo esta expresión de Korsgaard (1996), quien la utiliza, empero, en un sentido más limitado.

De manera más ambiciosa, responderé que el esclarecimiento tiene también su lugar en un debate político. ¿Es cierto, es coherente, decir que el bien común sólo podría realizarse a través de la búsqueda de los bienes particulares? ¿Es verdad que cuanto más racionales son los actores, mejor se satisfacen las exigencias de la razón? ¿O, a la inversa, hay que ver la racionalidad de los individuos como un obstáculo a la razón? Piénsese, por ejemplo, en la «paradoja del voto», resultante del hecho de que el actor racional no tiene ninguna razón para votar.[5] En efecto: la posibilidad de tener alguna influencia sobre el resultado del escrutinio es mucho menor que el riesgo de morir en un accidente de tránsito mientras se acude a las urnas. Además, quienes están en mejores condiciones de comprender la lógica de este razonamiento —en especial, los economistas de profesión— escogen con menor frecuencia la estrategia cooperativa en el dilema del prisionero, del cual el voto es un ejemplo clásico.[6]

[5] Véase Blais (2000).

[6] Frank, Gilovich y Regan (1993).

En tanto que la teoría de la elección racional ha sido objeto de amplias y muy precisas elaboraciones, no sucede lo mismo con la idea de la razón. La concepción que voy a proponer aquí no se apoya en una definición canónica, puesto que no la hay. Representa una síntesis personal —aun cuando no demasiado idiosincrásica, espero— de los textos clásicos.

Comencemos con una observación de La Bruyère: «No pensar más que en sí mismo y en el presente, fuente de error en la política».[7] Para corregir este error es necesario considerar a los otros y al futuro. De manera más precisa, hay que sustituir las perspectivas parciales que constituyen el egoísmo y la miopía por una actitud imparcial.

La idea de que la razón exige un tratamiento imparcial de los individuos corresponde a principios muy conocidos. Para resolver las cuestiones de justicia distributiva, Leibniz nos propone la siguiente máxima: «Poneos en el lugar de todos».[8] En las teorías recientes, esto equivale a decir que la elección de una organización justa de la sociedad

[7] La Bruyère (1990), pág. 377.

[8] Véase Elster (1975), págs. 127-30.

debe efectuarse detrás de un «velo de ignorancia», idea que puede interpretarse de varias maneras.[9] Para el utilitarismo, cada uno debe valer uno, y ninguno, más de uno. Para John Rawls, se debe escoger la forma de sociedad que beneficie a los menos favorecidos, sean quienes fueren. Otra idea imparcial es la de los derechos universales, encarnada en dos declaraciones, la de 1776 y la de 1789.

No se ha insistido tanto en la idea, de similar importancia, de que la razón requiere un tratamiento imparcial de los momentos temporales. En sí misma, ninguna fecha podría imponer un privilegio. Tomemos en primer lugar un ejemplo absurdo:[10] preferir siempre los productos que llegan un jueves a los que llegan un miércoles, tan sólo a causa de una predilección por ese día de la semana. Como lo vamos a ver, esta actitud no es contraria a los principios de la elección racional, pero sin duda es contraria a la razón. La mera preferencia por los jueves es una razón, pero la razón demanda también la razón de esta razón. Y evidentemente no la hay.

[9] Véase, por ejemplo, Fleurbaey (1996).

[10] Tomado de Parfit (1984), págs. 124-5.

Consideremos otro ejemplo, menos absurdo: preferir recibir cien euros hoy a recibir doscientos dentro de un año. Esta preferencia no es forzosamente contraria a la razón. Si mi esperanza de vida es de menos de un año o necesito los cien euros para sobrevivir, está enteramente fundada. Si es el simple resultado de la deficiencia de nuestra facultad telescópica, como dicen los economistas, es contraria a la razón. Desde el punto de vista objetivo, una persona que toma en cuenta las consecuencias remotas de las acciones presentes tiene más posibilidades de llevar una vida larga y feliz que aquella que sólo se interesa por los efectos inmediatos. Se verá que este hecho no tiene pertinencia alguna desde el punto de vista subjetivo.

Puede considerarse, en esta perspectiva, la idea del «interés bien entendido», tal como la utiliza, por ejemplo, Tocqueville. Una vez más, a falta de definiciones explícitas en los autores clásicos, hay que tratar de sintetizar sus ideas. Me parece que el interés bien entendido tiene al menos dos componentes. Por una parte, toma en cuenta las consecuencias remotas de la acción. En lenguaje técnico, corresponde a una elevada tasa de descuento

del futuro. Por la otra, se basa en creencias bien fundadas, en un sentido que voy a explicar.

Así caracterizado, el interés bien entendido es una amalgama de elementos objetivos y subjetivos. Una concepción íntegramente objetiva sustituiría las creencias bien fundadas por las creencias verdaderas. Empero, es imposible hacer que las decisiones políticas dependan de la posesión de la verdad. A lo sumo, se puede pedir que estén fundadas en creencias racionales, es decir, que sean resultado de un tratamiento no sesgado de una cantidad óptima de informaciones, para resumir una idea compleja que desarrollaré un poco más adelante. Baste con decir aquí que, dado que el óptimo de información depende de la tasa de descuento del futuro, las restricciones objetivas que pesan sobre esa tasa introducen también un elemento objetivo en el óptimo.

En esta reconstrucción, la idea de la razón implica tres elementos: la imparcialidad entre las personas, la imparcialidad temporal y las creencias racionales o bien fundadas. Habría que agregar sin duda la benevolencia, para excluir la malignidad imparcial. A veces nos topamos con la sugerencia

de que el primer elemento (la imparcialidad entre las personas) es redundante, ya que se deduce del segundo (la imparcialidad temporal). Desde la correspondencia de Descartes con la princesa Isabel hasta la teoría de los juegos repetidos, se ha señalado a menudo que el egoísmo previsor es capaz de remedar el altruismo.[11] Las condiciones en las cuales se puede contar con la actuación de esa mano invisible son, sin embargo, relativamente restrictivas.

La teoría de la elección racional es ante todo normativa y sólo secundariamente explicativa.[12] En primer lugar expone la manera en que los agentes tienen que actuar para alcanzar los fines que persiguen, y a continuación propone explicar sus acciones mediante la hipótesis de que los agentes se comportan efectivamente de esa manera. (Véase la figura 1.)

[11] Sobre este punto, véase Elster (2006*a*).

[12] Los economistas que creen que la racionalidad es pura coherencia podrían rechazar esta afirmación. Para convincentes refutaciones de este punto de vista, véase Sen (1973) y Hausman (1990).

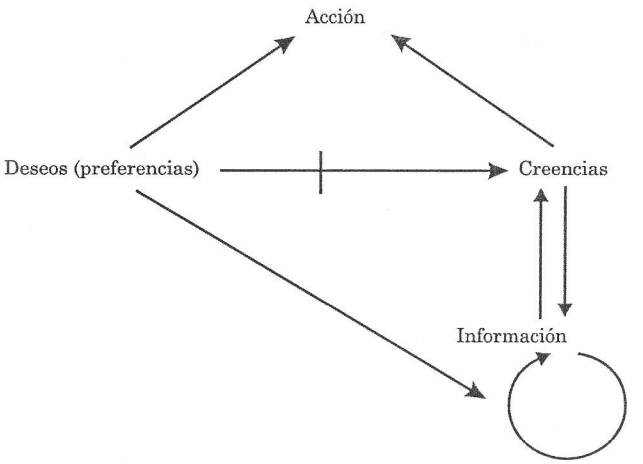


Figura 1.

Los deseos y las creencias son las razones de la acción. Un actor racional escoge la acción que realiza su deseo en la mayor medida posible, conforme a sus creencias y a la totalidad de sus demás deseos. Se trata de razones suficientes, que determinan de manera única lo que hay que hacer. Volveré a referirme a la posibilidad de no-unicidad, pero por el momento me atengo al caso ideal.

Los deseos comprenden, a la vez, las preferencias que podríamos calificar de sustanciales, como el hecho de preferir las manzanas a las naranjas, y las preferencias formales, como la actitud ante el

riesgo y el futuro. Se puede preferir la certeza de obtener cien euros a una lotería que prometa noventa o ciento veinte con una posibilidad sobre dos. También pueden preferirse cien euros hoy a ciento diez mañana. En condiciones precisas y bastante razonables, podemos representar esas preferencias mediante una *función de utilidad*, que asigne a cada opción un valor numérico. Esto permite entonces decir que el agente racional maximiza su utilidad.

Esta expresión no implica el egoísmo que en ocasiones se le atribuye. Cualquier deseo coherente, ya sea egoísta, altruista o malvado, es compatible con las exigencias de la racionalidad. Sólo quedan excluidos los deseos incoherentes, como, por ejemplo, el anhelo de que todo el mundo gane más que el promedio, o el de estar presente en los funerales de uno mismo para escuchar la oración fúnebre, a semejanza de Tom Sawyer y Huckleberry Finn. Esta última contradicción, el deseo de ser a la vez en-sí y para-sí, define según Sartre la existencia humana. Si es indudable que ese deseo puede generar acciones, estas no pertenecen al dominio de lo racional.

Un deseo puede también mostrar una incoherencia pragmática, en el sentido de que los medios utilizados para alcanzar el estado deseado obstaculizan su logro. Como escribió Paul Veyne, «sólo una expresión que no procure surtir efecto lo produce».[13] Sin embargo, antes de Veyne ya había dicho Proust: «Cada artista parece así como el ciudadano de una patria desconocida, olvidado de ella. (. . .) Los músicos no recuerdan esa patria perdida, pero cada uno de ellos se mantiene siempre inconscientemente armonizado en cierto unísono con ella; delira de alegría cuando canta al modo de su patria, la traiciona a veces por amor a la gloria, pero entonces, en busca de la gloria, huye de ella y únicamente cuando la desdeña la encuentra».[14]

Podríamos analizar de la misma manera las infructuosas tentativas de superar, mediante el solo esfuerzo de la voluntad, el insomnio, la impotencia sexual o la tartamudez. Esto no excluye que los

[13] Veyne (1976), pág. 679; véase también mi estudio sobre este libro en Elster (1990), capítulo 1.

[14] Proust (1988*b*), pág. 761; véase también, en el mismo sentido, *ibid.*, pág. 872.

estados deseados puedan alcanzarse de manera indirecta, como lo testimonia la existencia de somníferos y del Viagra. Siempre en la tradición sartreana, podemos citar también las conminaciones paradójicas, como «sé espontáneo».[15] En un registro religioso, las teologías enseñan a menudo que una acción emprendida con el exclusivo fin de ganar el Paraíso puede no dar acceso a él.[16] Cabría agregar que, desde el punto de vista moral, los intentos de algunos colaboracionistas franceses de redimirse, en 1944, mediante actos de resistencia cuyo único fin era ese no deberían haber terminado, como ocurrió a veces, en el sobreseimiento.[17]

Los deseos también corren el riesgo de ser incoherentes en un tercer sentido, si su estructura interna subvierte su realización. Para ilustrar esta idea, tomemos el ejemplo de la elección intertemporal.[18] De manera general, el valor presente de

[15] Véase, en especial, Watzlawick (1986).

[16] Véase, por ejemplo, Elster (2004*a*), para el caso de las misiones suicidas cumplidas con fines de salvación personal.

[17] En Elster (2006*b*) se hallarán ejemplos y análisis; véase, asimismo, Chauvy (2003).

[18] Véase Loewenstein y Elster (1992).

un bien futuro se puede representar como una función del tiempo que separa el presente de ese porvenir. En la concepción clásica se estipulaba un descuento exponencial del futuro, lo cual implica que las curvas correspondientes a dos bienes futuros distintos, uno pequeño e inmediato, otro más grande y más remoto, nunca se crucen. Sin embargo, según las investigaciones más recientes, parecería que ese descuento adopta, por lo común, una forma hiperbólica. Con el transcurso del tiempo, llega un momento en que el bien alejado, que en un principio parecía lo más deseable, deja de serlo, y esto lleva al agente a elegir el bien menor pero más cercano. (Véase la figura 2.)

Las creencias involucradas en ese cálculo se relacionan ya sea con los hechos particulares, ya con las relaciones causales. Por un lado, el agente no puede sino escoger entre las opciones que cree tener a su disposición. La existencia objetiva de una opción superior a aquellas de las cuales es consciente no podría influir en su acción. Por otro lado, el agente escoge entre las opciones de las que es consciente en función de las consecuencias posibles que les atribuye y de las probabilidades

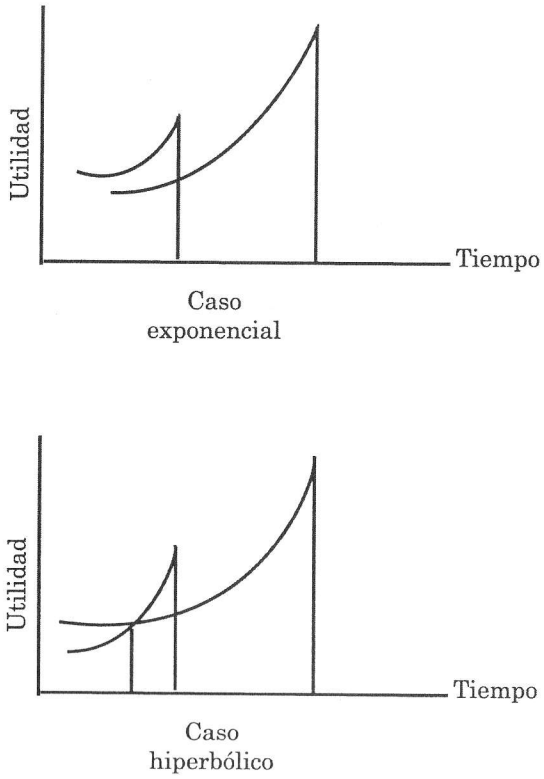


Figura 2.

que ellas tienen, a su entender, de producirse. De tal modo, la utilidad de las opciones se deduce de la utilidad de las consecuencias, ponderada por su probabilidad y reducida a un valor presente por la tasa de descuento del agente.

Para que la acción sea racional es menester que las creencias sobre las cuales se funda estén, a su

vez, bien fundadas. Esta exigencia se divide, por su parte, en dos. Por un lado, es indispensable que las creencias no estén desviadas con respecto a la información con que cuenta el agente. Aun cuando este pueda engañarse, no debe hacerlo de manera sistemática. Dado que las desviaciones son innumerables, sería deseable que se pudiera proponer una definición positiva, pero nadie ha conseguido hacerlo. Ocurre que en la formación de las creencias suele participar un elemento irreductible de *juicio*, o de apreciación subjetiva de la importancia relativa de los diversos elementos de información, con frecuencia muy heterogéneos entre sí, de que dispone el agente. Así, se ha llegado a asegurar que el presidente de la Reserva Federal, Alan Greenspan, tenía un conocimiento casi intuitivo de los mercados, sin par entre los teóricos de la economía.[19]

Las desviaciones son o bien «calientes», esto es, producidas por el sistema motivacional del agente,

[19] Paul Krugman, «Bernanke and the bubble», *The New York Times*, 28 de octubre de 2005. Cabría señalar que esta opinión de Krugman sobre Greenspan se ha modificado desde entonces.

o bien «frías», más similares a las ilusiones ópticas. Las primeras, a las cuales podemos darles el nombre de «creencias motivadas», que se indican mediante la flecha barrada en la figura 1, se conocen desde siempre. Se las resume en numerosos dichos, como «Cada quien cree con facilidad tanto en aquello que teme como en aquello que desea». La tendencia a creer en lo que se teme muestra que las creencias motivadas no tienen necesariamente la forma de «tomar los deseos por realidades». Compárese entre sí, por ejemplo, a esos dos maridos celosos que son el señor De Rênal y Otelo: el primero cree en lo que desea, y el segundo, en lo que teme.[20]

Las desviaciones frías tienen también una larga historia. Así explica Montaigne, por ejemplo, los errores de razonamiento que inducen a la gente a creer en la exactitud de las adivinaciones: «Así respondió Diágoras (. . .) a aquel que, al mostrarle en el templo muchos exvotos y cuadros de los que

[20] El señor De Rênal es uno de los personajes de *Rojo y negro*, la novela de Stendhal. Véase en Mele (2001) uno de los escasos tratamientos del tema que toman en cuenta tanto las creencias «contramotivadas» como las motivadas.

habíanse salvado del naufragio, le dijo: "¡Y bien! Vos, que pensáis que los dioses no se ocupan de las cosas humanas, ¿qué decís de tantos hombres salvados por su intercesión?". "Ocurre", respondió él, "que no están representados aquellos que se ahogaron, que son muchos más"».[21] La psicología moderna ha identificado una gran cantidad de errores similares que, por razones obvias, se prestan más a la experimentación que los errores calientes.

Para un actor racional, la información de la que se vale para formar sus creencias es más una variable que un dato. Antes de decidir cómo va a actuar, tiene que tomar una decisión acerca de la cantidad de recursos que invertirá en la búsqueda de los hechos pertinentes. El general o el médico sólo deben atacar u operar luego de un examen del terreno o del paciente. Sin embargo, es preciso que no se demoren demasiado en esa tarea, a riesgo de ser sorprendidos por el ataque o la huida del enemigo o por la muerte del paciente. Hay, pues,

[21] Montaigne (1991), *Ensayos*, I, 11; se hallará una compilación de estudios recientes en Gilovich, Griffin y Kahneman (2002).

en principio, una inversión óptima, que depende de las preferencias del agente, así como de sus creencias en lo tocante a los costos y los beneficios esperados. En la compra de un automóvil, por ejemplo, un agente racional que atribuye poca importancia al futuro remoto no hará un gran esfuerzo por comparar la durabilidad de los distintos modelos. Tampoco consultará a su médico para interiorizarse más sobre los riesgos que puede traerle aparejados el consumo de *foie gras*, alcohol o cigarrillos.

Para poner de relieve la naturaleza radicalmente subjetiva de la noción de racionalidad que acabo de exponer, consideremos un ejemplo hipotético. Pensemos en una persona que sufre —la palabra no puede ser más adecuada— una tasa de descuento muy baja, es decir, que adjudica escasa importancia a las consecuencias remotas de su acción. A raíz de su imprevisión, se derrumban sus finanzas y su salud. Y lo que se le propone como remedio para su problema es. . . una píldora que le hará atribuir mayor importancia al futuro. ¿La tomará? En el caso de una persona racional, la respuesta debe ser negativa. Todas las acciones que la

píldora podría inducirle ya forman parte de su repertorio. Si la persona no las elige es porque no desea llevarlas a cabo. No desea, por lo tanto, hacer la elección que la impulsaría a realizarlas.

La premisa de esta afirmación es lo que podríamos denominar *principio de no-indirección*. Llamo *indirección* a una operación indirecta que permite llegar a cierto resultado por medio de dos acciones sucesivas, la primera de las cuales sólo sirve para hacer posible la segunda. Por ejemplo, para encontrar mis llaves debo ante todo ponerme los anteojos. En este caso, no habría ninguna objeción a encontrarlas sin ponérmelos; el único inconveniente residiría en que no podría. En el caso de la píldora, la situación es inversa: podría comenzar desde ya a velar por mis finanzas y mi salud; el único inconveniente es que no quiero. El principio de no-indirección afirma que un agente racional se niega a hacer en dos pasos lo que no querría hacer en uno, lo cual no excluye las operaciones indirectas en procura de los fines que él no *puede* alcanzar por medio de una sola acción.

Este principio no sólo rige las acciones racionales. Consideremos otro ejemplo hipotético. Tras

perder la cabeza, me he enredado en una relación adúltera. Mi esposa no sabe nada y, supongámoslo, jamás lo sabrá. Tengo, empero, un sentimiento de culpa muy grande que me impulsa a romper el amorío. Supongamos ahora que un amigo me ofrece una píldora capaz de suprimir ese sentimiento, lo cual me permitiría continuar con la relación sin problemas de conciencia. El principio de no-indirección prescribe que debo rechazar la píldora, porque me sentiría tan culpable al aceptarla como al seguir adelante con la relación. Mi amigo tiene otra opinión: estima que, una vez ingerida la píldora, yo estaría más contento. Yo mismo, tras tomarla, estaría feliz de haber procedido así, pero esto no afecta en absoluto el hecho de que, desde mi punto de vista *ex ante*, el rechazo se imponga. Si se acepta esta conclusión, también hay que rechazar la idea de que las emociones sólo intervienen en las decisiones en virtud de los placeres o los dolores asociados.[22] Es totalmente racional, antes de una fiesta, tomar una píldora que

[22] Idea que se desarrolla, sobre todo, en los trabajos de Gary Becker; en Elster (1999), capítulo 4, se hallará un tratamiento crítico.

a la mañana siguiente me evitará una resaca; pero un sentimiento de culpa no es lo mismo que un dolor de cabeza.

En consecuencia, el principio de no-indirección expresa una idea de coherencia más amplia, que abarca tanto lo racional como lo emocional. No obstante, las violaciones del principio son bastante frecuentes, como en el caso de esos empleados alemanes y noruegos que son demasiado orgullosos como para aceptar subsidios públicos a sus salarios, pero que toleran que su empresa pague una tarifa eléctrica reducida con el solo objetivo de mantener los puestos de trabajo.[23] Otro ejemplo de violación del principio: en la Convención Federal de Filadelfia de 1787, James Wilson propuso, para calmar a quienes se oponían a la idea de incluir a los esclavos en la base de la representación electoral, que se los incorporase simplemente en la base de la imposición y, a continuación, se hiciera de esta la base de aquella.[24]

También puede expresarse la subjetividad de lo racional diciendo que el ojo no puede ver más allá

[23] Véanse Schlicht (1984) y Serck-Hanssen (1971).

[24] Farrand (1966), vol. 1, pág. 561.

de su horizonte. Para valernos otra vez del ejemplo de las llaves y los anteojos, modifiquemos la situación y consideremos que se han perdido estos últimos; por lo tanto, no puedo ponérmelos para encontrarlos. Otra metáfora: el agente racional está preso en la trampa de sus deseos y sus creencias. Acabamos de ver en qué sentido puede esto ser cierto en cuanto a la actitud con respecto al futuro. Para mostrar que la idea se aplica igualmente a las creencias, puede definirse como *creencia-trampa* a aquella cuyos costos previstos de verificación son tan elevados que un agente racional tiene que renunciar a verificarla.[25] En Polonia, hasta muy poco tiempo atrás, muchas personas que libraban una dura batalla contra su alcoholismo se hacían implantar el medicamento disulfiram bajo la piel.[26] En vista de que al ser administrado por vía oral el disulfiram provoca, como único efecto, náuseas violentas cuando se bebe alcohol, se suponía que su aplicación subcutánea causaba la muerte en caso de ingesta alcohólica. A decir verdad, esta sustancia no tiene ningún efecto

[25] Mackie (1996).

[26] Osiatynski (1997).

cuando se la implanta,[27] pero, en tanto la gente creyera que lo tenía, era racional de su parte no tratar de comprobar la veracidad de esta creencia.

La teoría de la acción racional toma como un dato fijo los deseos y las preferencias del agente. En la figura 1, ninguna flecha apunta hacia los deseos. En la tradición de Hume, los deseos son las causas primeras de la acción. Uno no escoge sus deseos. Hay, claro está, contraejemplos de esta proposición. Podemos someternos a una psicoterapia para reducir la fuerza de algunos deseos si nos vemos en la imposibilidad de oponerles resistencia. Ahora bien, este caso presupone que el agente es irracional, porque es víctima de una debilidad de la voluntad. Al escoger la acción que realiza su deseo, no tiene en cuenta la totalidad de sus deseos. Para un actor racional, el problema de escoger sus deseos no se plantea. Desear tener cierto deseo es ya tenerlo. Querer ser movido por las consecuencias remotas de la acción es ya tener esa motivación.[28]

[27] Johnsen y Mørland (1992).

[28] Sin embargo, la cuestión es más complicada; véase Elster (2007).

La teoría de la acción racional adolece de lagunas, pues en algunos casos es imposible determinar mediante ella una acción única que derive de los deseos y las creencias del agente. Desde el punto de vista de la sociología de las ciencias, quizá sea la promesa de unicidad la que explique en gran parte el éxito fulminante de la teoría. Admitámoslo: es un hecho matemático elemental que una función calificada de «bien ordenada», que se define sobre la base de un conjunto «bien ordenado», alcanza su máximo para un valor único de la variable independiente. La determinación de la elección del consumidor como el punto de tangencia entre la línea que representa la restricción presupuestaria y una curva de indiferencia propone una ilustración clásica. Como las ecuaciones newtonianas, el modelo del agente racional parece encarnar, por ende, el ideal científico de la predicción única.

Sin embargo, desde hace tiempo se sabe, en física, que el problema de *n* cuerpos no tiene, en el caso general, ninguna solución analítica. De manera vagamente parecida —se trata de una comparación pedagógica, y nada más—, hay situaciones en

que la interacción de *n* agentes no tiene solución óptima o equilibrio único. No es mi intención examinar aquí este aspecto de la teoría de los juegos. Prefiero insistir en otra fuente de no-unicidad de la acción racional, a saber: la adquisición de información.

Supongan que han ido a recolectar fresas silvestres, tan deliciosas como escasas y difíciles de encontrar.[29] Están de vacaciones en una región que no conocen muy bien, y tienen muy pocas ideas preconcebidas sobre la distribución de esos lugares secretos y mágicos que dan título a la película de Bergman, *Smultronstället* (mal traducida al francés como *Les Fraises sauvages*).* Buscan un poco al azar en los prados, sabiendo que en todos ellos van a encontrar al menos unas cuantas fresas dispersas, pero que algunos les permitirían llenar su canasta en una hora. La pregunta que se plantea

[29] Ejemplo tomado de Johansen (1977), pág. 144; véase también Winter (1964), pág. 262.

* El filme se conoció en la Argentina con el título de *Cuando huye el día*, y como *Fresas salvajes* en España. Literalmente, el título sueco se podría traducir aproximadamente como «Un lugar hermoso y preferido». (*N. del T.*)

es: ¿Cuándo dejar de buscar y empezar, mal o bien, a recolectar?

Dado que sólo disponen de un día, ustedes saben que tendrán que dejar de buscar indefectiblemente antes de que caiga la noche. Sin embargo, también saben que comenzar la recolección en el primer prado que encuentren no es, por lo general, la estrategia óptima, salvo, desde luego, que se trate de uno de esos lugares mágicos, eventualidad representada por el rizo de la figura 1. Entre el máximo y el mínimo de inversión hay, sin duda, un óptimo objetivo, pero, ¿cómo determinarlo? A menudo, el general y el médico mencionados algunos párrafos más atrás se hallan en la misma situación. Entre el mínimo y el máximo de tiempo que deben consagrar a la recopilación de información hay una zona de indeterminación, que puede ser más o menos grande según los casos, pero que suele ser considerable. Aun cuando un buen juicio permita formarse una creencia racional a partir de una cantidad de información dada, a veces es incapaz de determinar la cantidad óptima.

Por lo tanto, la teoría de la acción racional puede fracasar, porque es inhábil para generar pres-

cripciones y predicciones únicas. También puede fracasar si los agentes no ajustan su comportamiento a las predicciones, sean estas únicas o no; es decir, si ellos son irracionales. Hay múltiples fuentes de irracionalidad, calientes o frías. En lo que aquí me interesa, querría mencionar sobre todo el papel de las emociones, que son capaces de actuar en cada uno de los cuatro polos de la acción:

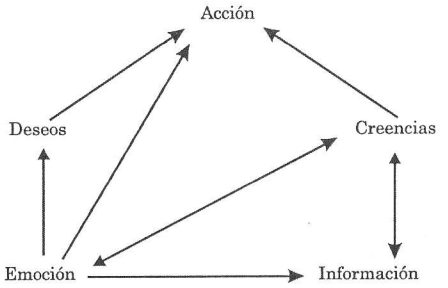


Figura 3.

La debilidad de la voluntad —actuar contra la totalidad de los deseos propios bajo la influencia de uno de ellos— se debe, en muchos casos, a que el deseo dominante surge de una emoción fuerte. Como dice la Medea de Ovidio cuando va a matar a sus hijos para vengarse del padre de estos, que la ha engañado y abandonado: «Veo el bien y lo apruebo, y a lo peor me entrego».

Mediante otro mecanismo, las emociones efectúan una inversión o un vuelco temporal de las preferencias. Una ilustración simple de esto es la deserción frente al peligro. Uno puede tener la inconmovible intención de resistir a pie firme el ataque enemigo, pero abandonarla bajo la conmoción del terror y lamentar amargamente ese abandono. En el ejército, el remordimiento suele inducir a los desertores a regresar. Así, en la Guerra de Secesión norteamericana, el diez por ciento de los doscientos mil hombres que habían desertado del ejército de la Unión retornaron a él por voluntad propia.[30]

Así como las emociones nacen de las creencias, influyen a su vez en estas. Por un lado, nos dice Stendhal, «basta con una muy pequeña medida de esperanza para que nazca el amor». Por el otro, «desde el momento en que ama, el hombre más sabio ya no ve ningún objeto *tal como es*. (. . .) Ya no atribuye nada al azar; pierde el sentido de la probabilidad; una cosa imaginada es una cosa que existe por el efecto sobre su felicidad».[31]

[30] Costa y Kahn (2007).

[31] Stendhal (1980), capítulos 3 y 12.

Por último, la *urgencia* de las emociones induce una preferencia por la acción inmediata en desmedro de la acción diferida. Distingo la urgencia de la impaciencia, que por su parte induce una preferencia por una recompensa inmediata en comparación con una recompensa diferida. En esta acepción del término, la urgencia produce un cortocircuito en la adquisición de la información. Séneca alababa a Fabio por saber «esperar, contemporizar, cosas de las cuales el hombre irritado es incapaz». De manera más general, agrega, «la razón confiere a cada parte el lugar y el tiempo convenientes; ella misma se impone plazos para disfrutar de toda libertad en la discusión de la verdad. La cólera lo hace todo corriendo».

No obstante, es indudable que hay situaciones peligrosas en las cuales se impone la acción inmediata, pues en esos casos la espera sería desastrosa. La dificultad consiste en distinguir con claridad entre los sobresaltos justificados y las reacciones de pánico. Luego del 11 de septiembre de 2001, ¿fueron apropiadas o excesivas las reacciones de los gobiernos occidentales? ¿Atenuaron el peligro o lo hicieron mayor aún? La cuestión no estriba

sólo en saber si la deliberación fue insuficiente, sino también si fue sesgada en el sentido de que se cree muy fácilmente en lo que se teme.[32]

Me gustaría ahora ampliar el análisis y distinguir entre deseos y motivaciones. En tanto que los primeros se refieren directamente a la acción a emprender, las segundas son las actitudes más fundamentales que generan los deseos. Así, el deseo de castigar a quien nos ha ofendido puede nacer ya sea de un cálculo interesado, ya sea de una emoción, o bien, por último, de un principio imparcial de justicia retributiva. Como señalé al comienzo y como lo sugiere este ejemplo, seguiré a los moralistas y distinguiré tres motivaciones principales: la razón, las pasiones y el interés.

Ya he procurado explicar en qué consiste la razón. Voy a considerarla ahora en cuanto motivación interiorizada por el agente, en un sentido un tanto indirecto, como vamos a ver. Concibo a la pasión en un sentido lato, que abarca no sólo las emociones sino también la locura, la embriaguez y

[32] Sobre este punto, véase Elster (inédito).

los apetitos de los toxicómanos. El interés, finalmente, se comprende en el sentido de la ventaja personal, ya sea por los bienes materiales, el poder, el prestigio o la gloria.

En toda sociedad existe una jerarquía normativa de las motivaciones. Se nos alaba por haber cumplido una acción dada con tal o cual motivación, independientemente del elogio o de la censura asociados a la acción misma. En la antigua Grecia, el patriotismo, la ambición de defender a la Ciudad, estaba en la cima de la jerarquía; posteriormente se ubicaba la venganza apasionada a causa de una afrenta personal, y luego el interés, el deseo sexual y la ebriedad. En el fondo de la escala estaban la envidia y la *hybris* (o desmesura). En otras sociedades, la venganza precede al patriotismo, e incluso en algunas se sitúa por debajo del interés en la escala de valores.

Según Tocqueville, los norteamericanos de su tiempo invertían el orden entre la razón y el interés en la jerarquía: «[A diferencia de Europa, donde] aún se fingen todos los días (. . .) grandes devociones que ya no existen, los norteamericanos, al contrario, consienten en explicar, por medio del

interés bien entendido, casi todos los actos de su vida; muestran con complacencia que el amor iluminado por sí mismos los lleva sin cesar a ayudarse unos a otros y los dispone a sacrificar de buena gana, para el bien del Estado, una parte de su tiempo y de sus riquezas. Creo que en este aspecto suelen no hacerse justicia; en efecto: a veces, tanto en Estados Unidos como en otros lugares se ve a los ciudadanos abandonarse a los impulsos desinteresados e irreflexivos que son naturales en el hombre; pero los norteamericanos casi nunca confiesan que ceden a mociones de esta suerte: prefieren hacer honor a su filosofía antes que a sí mismos».[33]

Estudios recientes sugieren que esta actitud persiste bajo diversas formas en las sociedades occidentales.[34] La cuestión se complica debido a que la fuente de desaprobación de las motivaciones imparciales o altruistas es, con frecuencia, la sospecha de que el presunto desinterés no es, en los hechos, más que una pose. Para disipar esa sospecha, incluso una persona angelical como la abuela

[33] Tocqueville (1992), págs. 636-7.

[34] Miller (1999).

proustiana atribuía «siempre móviles egoístas a lo que hacía».[35] Aunque en la jerarquía la razón domine al interés, este último premia la afectación ostentosa de razón. Para citar a Montaigne: «Cuanto más llamativa es una buena acción, más descuento de su bondad la sospecha que en mí nace de que haya sido realizada más por ser llamativa que por ser buena; al ser exhibida, casi se vende».[36] Vamos a ver que Montaigne podría haber sido aún más radical.

Demos por establecida la existencia de una jerarquía de las motivaciones, sea cual fuere, y tratemos de determinar sus efectos. Para granjearse alabanzas o evitar reprobaciones, el agente racional está interesado en presentar su motivación como si estuviera situada en un lugar de la jerarquía más elevado que el que realmente le corresponde. Puede tener interés en la apariencia de la razón e incluso en la apariencia de la emoción, como en las sociedades en que la búsqueda de una esposa rica es un móvil inconfesable. Sin embargo, al mismo

[35] Proust (1988*a*), pág. 28.

[36] Montaigne (1991), *Ensayos*, III, 10.

tiempo que mantiene esa motivación de segundo orden, el agente conserva su deseo inspirado en una motivación de primer orden, y así se plantea un potencial conflicto.

En la hipocresía, se trata de engañar a los otros sobre las verdaderas motivaciones que uno abriga. La acción ante los dicasterios atenienses ofrece numerosos ejemplos de ello. Así, los acusadores que temían ser acusados ellos mismos de sicofantes se preocupaban por mostrarse motivados por la venganza, y no por su beneficio material.[37] En los debates públicos, toda propuesta interesada debe presentarse como relevante para el interés público. Los grandes partidos políticos, que están interesados en contar con un sistema electoral mayoritario, lo defienden en general en nombre de la eficacia; los partidos pequeños, interesados en un sistema proporcional, lo justifican, las más de las veces, en nombre de la democracia. Cuando hay una jerarquía interna en la razón, como cuando la utilidad social está subordinada a los derechos de propiedad, las propuestas motivadas por aquella

[37] Hansen (1991), pág. 195.

pueden defenderse en nombre de estos. Así sucedió con la confiscación de los bienes del clero en octubre de 1789.[38]

Sin embargo, el homenaje que el vicio le rinde a la virtud puede también adoptar la forma de la mala fe o del autoengaño. Según Jean Domat, «toda la deferencia que el corazón tiene por el espíritu radica en que, si no actúa por razón, al menos hace creer que actúa por razón».[39] La autoestima puede ser un móvil tan importante como la estima de los otros. De hecho, el propio acto de renunciar al público exterior suele ser aplaudido por el público interior, y puede llevarse a cabo en función de esos aplausos, más que por sí mismo. Como dice La Rochefoucauld, «el orgullo se resarce siempre y no pierde nada ni siquiera cuando renuncia a la vanidad».[40] Una donación caritativa, aun hecha en secreto, puede ser compensada por el refuerzo de amor propio que produce. En este aspecto, la sospecha de Montaigne no es todo lo radical que debería.

[38] Véase, por ejemplo, *AP*, 9, págs. 639 y sigs. y 649 y sigs.

[39] Domat (1992), pág. 611.

[40] La Rochefoucauld (1982), máxima 33.

A menudo nos esforzamos, empero, por *justificar* mediante la razón las acciones interesadas, en vez de sacrificarlas por ella. Aunque mi interés me aconseje no donar nada a las buenas obras, quiero conservar la imagen de mí mismo como la de una persona que no es movida tan sólo por su interés. Puedo satisfacer ambas exigencias si adopto un principio de caridad que sea apropiado a las circunstancias. Si otros donan mucho, yo me abstengo diciéndome que una contribución de mi parte produciría escaso bien adicional.[41] Si otros donan poco, yo vuelvo a abstenerme diciéndome que la equidad sólo exige que sea generoso si los demás también lo son.[42] Tanto en un caso como en el otro, me quedo con el oro y con el moro. De manera similar, como dice La Rochefoucauld, «el rechazo de las alabanzas es un deseo de ser alabado dos veces».[43]

Las pasiones, asimismo, buscan con frecuencia el aval de la razón. Como afirma Séneca, «la razón

[41] Margolis (1982).

[42] Sugden (1984).

[43] La Rochefoucauld (1982), máxima 149.

quiere decidir lo que es justo; la cólera quiere que parezca justo lo que ha decidido».[44] En 1945, los franceses, como los belgas, los noruegos, los daneses, los húngaros o los holandeses, estaban desgarrados entre dos móviles: el deseo apasionado de venganza y el deseo de respetar los principios del derecho, sobre todo el de no retroactividad de las leyes. En muchos casos, de hecho, los actos que se quería castigar no estaban prohibidos por las leyes que se hallaban vigentes en el momento en que se los había cometido.

Los húngaros, los daneses y los holandeses se inclinaron decididamente por la retroactividad. En Francia, Bélgica y Noruega, leyes que eran retroactivas fueron presentadas como si no lo fueran. En el caso francés se dijo que, en vista de que los actos calificados de indignidad nacional eran de orden político, y no criminal, no había violación del principio que reza: «Ningún delito sin ley». No se trataba más que de un juego de palabras, porque la degradación nacional que se aplicó como castigo por esos actos constituía una viola-

[44] Séneca, *De la cólera*, I, 18.

ción flagrante de otro principio, de igual importancia: «Ninguna pena sin ley».[45]

Una pasión puede también buscar el aval de otra situada en un lugar más alto de la jerarquía. Como señala Plutarco, los hombres, con frecuencia, presentan la emoción inconfesable de la envidia bajo el disfraz de la ira o del odio.[46] El antisemitismo, por ejemplo, suele tener origen en una envidia transmutada en indignación autosatisfecha. En otro registro, la iconoclasia de los hugonotes estaba a menudo motivada por un espíritu de odio y de venganza, más que por el amor a Dios.

Si bien tienen un campo de actuación muy vasto, estos mecanismos también están sometidos a restricciones importantes. Cierto es que su aplicación se ve facilitada por el hecho de que los agentes tienen *dos grados de libertad* en la armonización de sus motivaciones y sus deseos. En primer lugar, hay un gran número de concepciones de la imparcialidad: unas apelan a los derechos in-

[45] Sobre las observaciones de los dos últimos párrafos, véase Elster (2004*b*), págs. 237-40.

[46] Plutarco, «Sobre la envidia y el odio».

dividuales; otras, a la noción de equidad, y otras, incluso, a la elección detrás del velo de ignorancia. Para un mismo caso, las implicaciones de la acción pueden ser muy diversas, como acabamos de verlo en lo referido a la caridad.

En segundo lugar, toda decisión política de cierta importancia está imbricada en cuestiones de causalidad social respecto de las cuales pueden defenderse, con algún grado de verosimilitud, posiciones muy distintas. Ya se trate de los efectos del salario mínimo, del subsidio a los desocupados o de la tasa de imposición, siempre habrá profesores de economía que defiendan ideas completamente opuestas entre sí. Así, un agente tendría que ser muy inepto o muy desafortunado para no hallar una combinación de principios normativos y encadenamientos causales que le permitiera presentar su pasión o su interés particulares bajo una luz imparcial.

Sin embargo, estos mecanismos están sujetos a dos restricciones que hacen que su aplicación no sea tan sencilla. Ante todo, hay una restricción ligada a la coherencia: una vez que el agente adopta determinado principio normativo o cierta teoría

causal, no puede abandonarlos aunque ya no le permitan satisfacer sus deseos. Si he adoptado un principio de equidad para justificar mis contribuciones poco generosas, estoy obligado a persistir en él si los demás comienzan a hacer mayores donaciones, so pena de revelarme a mí mismo como alguien a quien sólo mueve su interés —precisamente, la apariencia que quería evitar—. Es más fácil entrar en una situación por mala fe que salir de ella por el mismo mecanismo.

Por otro lado, está lo que podemos llamar «restricción de imperfección», originada en la necesidad de evitar que la coincidencia entre la motivación profesada y el deseo sea demasiado flagrante. A fin de ocultar sus verdaderas motivaciones a uno u otro público, interior o exterior, uno debe con frecuencia, y en cierta medida, actuar en contra de sus deseos. Supongamos que quiero hacer creer que una propuesta de matrimonio está animada por el amor y no por el afán de lucro. Para sostener tal ilusión tengo que evitar que mi elección recaiga en una persona cuya única cualidad consista en poseer una gran fortuna, aun cuando sea esta, y sólo esta, la que dicte mi preferencia.

No obstante, frente a un público, exterior o interior, que lleva muy lejos la hermenéutica de la sospecha, la única manera de convencer consiste, a veces, en renunciar a los propios deseos. Pascal señala que «nuestro propio interés es también un maravilloso instrumento para arrancarnos los ojos agradablemente. No le está permitido al más ecuánime de los hombres ser juez en su propia causa. Sé de quienes, para no caer en ese amor propio, fueron los más injustos del mundo a contracorriente».[47] En ocasiones, sólo una actitud contraria al interés puede generar la apariencia del desinterés.

Esta deferencia para con la razón —para retomar las palabras de Jean Domat— tiene su contrapartida en cierta deferencia hacia la racionalidad. La racionalidad es una norma, por razones que competen a la lógica y no a una jerarquía social de valores. Si se quiere cumplir un deseo, no se puede no querer también hacerlo con el mínimo costo y evitando los gastos inútiles. Todo posible contraejemplo equivaldría a especificar el deseo inicial de una manera más rica. Por eso, la racionalidad, a

[47] Pascal (1995), pensamiento 78 (edición Sellier).

diferencia de la imparcialidad, es un valor transhistórico y transcultural. Las tentativas, numerosas, de deducir la primera de la segunda están condenadas al fracaso.

La norma de la racionalidad está en el corazón de la antropología filosófica, mientras que la norma de la imparcialidad concierne a la antropología empírica. En algunas sociedades, la negligencia o la violación de los valores imparciales no dan lugar a sanciones sociales. Esto no quiere decir que esas sociedades no conozcan los comportamientos altruistas o desinteresados, sino tan sólo que se los presenta —como señala Tocqueville acerca de los norteamericanos de su tiempo— como el fruto del interés bien entendido. La capacidad del egoísmo previsor de remedar al altruismo también le permite a este darse una apariencia egoísta.

Recordemos que la norma de la imparcialidad, en cuanto norma social, se refiere a la motivación de la acción, más que a la acción misma. Así pues, la norma es capaz, asimismo, de modificar la acción, como se deduce de inmediato de las restricciones de coherencia y de imperfección. Desde el punto de vista social, el efecto de esas restricciones

suele ser benigno. Se puede hablar no sólo de la fuerza civilizadora de la hipocresía, sino incluso de la fuerza civilizadora de la mala fe. La norma de la imparcialidad también puede ser objeto, no obstante, de un uso *estratégico* cuyo efecto es más ambiguo.[48] Aun en los contextos en que las motivaciones interesadas son lícitas, como en las negociaciones salariales, las partes recurren a veces a los principios de equidad o de igualdad. Dado que renunciar a un principio constituye una concesión mayor que renunciar a un interés, e impone, por consiguiente, una mayor contraconcesión, a las partes les interesa presentarse como no movidas exclusivamente por su beneficio. Si ese comportamiento se generaliza, el acuerdo puede entonces resultar imposible.[49] La apelación a la razón subvierte la razón.

De manera parecida, pero por un mecanismo diferente, la apelación a la racionalidad puede subvertir la racionalidad. El ideal de la decisión ra-

[48] Sobre este aspecto, véase Elster (1995), con ejemplos tomados de los debates constituyentes norteamericano (1787) y francés (1789).

[49] Elster (1989*b*), capítulo 6.

cional es actuar conforme a razones suficientes, a la luz de las cuales se pone de relieve una acción óptima única. En algunos casos, empero, la búsqueda de la acción óptima no vale la pena, y el actor racional adoptará entonces un proceder distinto —por ejemplo, seguir la tradición o decidir a cara o cruz—. Se puede, en efecto, definir la *hiperracionalidad* como la búsqueda de la acción que habría sido óptima si se ignorasen los costos de esa misma búsqueda.[50]

Los costos de la decisión se incluyen en tres categorías. En primer lugar, hay que considerar el costo directo, que depende de la inversión en la adquisición de información. Si vamos de tienda en tienda para comprar un producto determinado al precio más bajo posible, hay que tener en cuenta el costo del taxi o el del pasaje del metro. Tenemos a continuación el costo de oportunidad, que es el valor de la mejor utilización alternativa del tiempo dedicado a la recolección de información. Aun cuando el precio del pasaje del metro sea inferior a la ganancia bruta esperada por la compra al precio

[50] Los análisis de Neurath (1913) son aún los mejores.

más bajo, el agente podría haber dedicado los treinta minutos del viaje a actividades que para él tienen más valor que la ganancia neta. Por último, está el costo de los subproductos del proceso de decisión.

Puede observarse la negligencia respecto de los dos primeros costos en la conducta de quienes veranean en el Rosellón. Son muchos los que van en auto a España para comprar cigarrillos a bajo precio, como si el combustible y el tiempo no les costaran nada, por no mencionar las colas en los restaurantes de las autopistas, donde esperan media hora para engullir la (mediocre) comida que necesitan luego de tres horas de embotellamiento. Al regresar pueden felicitarse, desde luego, por haber ahorrado, pero a un precio absurdo.[51]

La negligencia respecto del tercer tipo de costo puede traer aparejados efectos más importantes, como en el caso de atribución del derecho de guarda de los hijos de una pareja a raíz de un divorcio.[52] En la mayoría de los países occidentales, la

[51] *L'Indépendant*, Perpiñán, 13 de agosto de 2005, pág. 2.

[52] Ejemplo desarrollado en Elster (1989*a*), capítulo 3.

ley prescribe que, en caso de conflicto entre los padres, la atribución debe respetar el «interés superior del niño», es decir, que la guarda debe corresponderle al progenitor más apto para promover el bienestar del hijo. Ahora bien, la experiencia demuestra ampliamente que la determinación jurídico-psicológica de la aptitud relativa de los padres es un proceso muy penoso y nocivo para el propio niño cuyo interés se busca supuestamente proteger. El interés del niño sería mejor respetado, sin duda, mediante la tradicional presunción en favor de la madre, o incluso, por qué no, mediante un sorteo entre ambos progenitores.

Las sociedades occidentales modernas están impregnadas de hiperracionalidad. En ese sentido, y sólo en ese sentido, puede hablarse de la racionalidad, o, mejor, de su abuso, como un fenómeno específicamente occidental o moderno. Se ha hablado, a este respecto, del efecto iatrogénico de la búsqueda de soluciones óptimas.[53] En cierta medida, puede decirse que «quien quiere lo más, hace lo menos».

[53] Wiener (1998).

¿Cuáles son, en definitiva, las funciones de la razón y de la racionalidad en los comportamientos humanos? Son las funciones del preceptor y del consejero del príncipe, respectivamente. El preceptor le enseña al príncipe a promover el bien público a largo plazo. El consejero le explica cómo actuar para realizar con la mayor eficacia sus fines, sean cuales fueren. No le incumbe al consejero imponer las exigencias de la razón, pero, si el preceptor ha hecho bien su trabajo, el príncipe las hará suyas.

Referencias bibliográficas

AP = Archives Parlementaires de 1787 à 1860, primera serie (1787 a 1799), París, 1875-1888.

Blais, André (2000) *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory*, Pittsburgh: University of Pittsburgh Press.

Chauvy, Gérard (2003) *Les Acquittés de Vichy: non-lieux et acquittements pour faits de résistance dans les procès de la Libération*, París: Perrin.

Costa, Dora L. y Matthew E. Kahn (2007) «Deserters, social norms, and migration», *The Journal of Law and Economics*, 50, págs. 323-53.

Damasio, Antonio (2000) *L'Erreur de Descartes: la raison des émotions*, París: Odile Jacob [*El error de Descartes: la emoción, la razón y el cerebro humano*, Barcelona: Crítica, 2006].

Domat, Jean (1992) «Pensées», en Jean Lafond (ed.), *Moralistes du XVII^e siècle*, París: Robert Laffont.

Elster, Jon (1975) *Leibniz et la formation de l'esprit capitaliste*, París: Aubier-Montaigne.

(1989*a*) *Solomonic Judgements: Studies in the Limita-*

tion of Rationality, Cambridge (Inglaterra) y Nueva York: Cambridge University Press [*Juicios salomónicos: las limitaciones de la racionalidad como principio de decisión*, Barcelona: Gedisa, 1991].

(1989*b*) *The Cement of Society: A Study of Social Order*, Cambridge (Inglaterra) y Nueva York: Cambridge University Press [*El cemento de la sociedad: las paradojas del orden social*, Barcelona: Gedisa, 1997].

(1990) *Psychologie politique*, París: Éditions de Minuit [*Psicología política*, Barcelona: Gedisa, 1995].

(1995) «Strategic uses of argument», en Kenneth Arrow *et al.* (eds.), *Barriers to the Negotiated Resolution of Conflict*, Nueva York: Norton.

(1999) *Alchemies of the Mind: Rationality and the Emotions*, Cambridge (Inglaterra) y Nueva York: Cambridge University Press [*Alquimias de la mente: la racionalidad y las emociones*, Barcelona: Paidós, 2002].

(2004*a*) «Motivations and beliefs in suicide missions», en Diego Gambetta (ed.), *Making Sense of Suicide Missions*, Oxford y Nueva York: Oxford University Press [«Las motivaciones y creencias en las misiones suicidas», en Diego Gambetta (ed.), *El sentido de las misiones suicidas*, México: Fondo de Cultura Económica, 2009].

(2004*b*) *Closing the Books: Transitional Justice in Historical Perspective*, Cambridge (Inglaterra) y Nueva York: Cambridge University Press [*Rendición de cuentas: la justicia transicional en perspectiva histórica*, Buenos Aires: Katz, 2006].

(2006*a*) «Altruistic behavior and altruistic motivations», en Serge-Christophe Kolm y Jean Mercier Ythier (eds.), *Handbook of the Economics of Giving: Altruism and Reciprocity*, vol. 1, Ámsterdam: Elsevier.

(2006*b*) «Redemption for wrongdoing», *Journal of Conflict Resolution*, 50, págs. 324-38.

(2007) *Agir contre soi*, París: Odile Jacobs.

(inédito) «Fear, terror, and liberty: a conceptual framework», trabajo preparado para el seminario sobre «Terrorism and Civil Liberties», desarrollado en la Columbia University en la primavera de 2006.

Farrand, Max (ed.) (1966) *The Records of the Federal Convention of 1787*, New Haven: Yale University Press.

Fisher, Nicholas Ralph Edmund (1992) *Hybris: A Study in the Values of Honour and Shame in Ancient Greece*, Warminster: Aris and Phillips.

Fleurbaey, Marc (1996) *Théories économiques de la justice*, París: Economica.

Frank, Robert, Thomas Gilovich y Dennis T. Regan (1993) «Does studying economics inhibit cooperation?», *Journal of Economic Perspectives*, 7(2), págs. 159-71.

Gilovich, Thomas, Dale Griffin y Daniel Kahneman (eds.) (2002) *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge (Inglaterra) y Nueva York: Cambridge University Press.

Hansen, Mogens Herman (1991) *The Athenian Democracy in the Age of Demosthenes: Structure, Principles, and Ideology*, Oxford: Blackwell.

Hausman, Daniel M. (1990) «Revealed preference, belief, and game theory», *Economics and Philosophy*, 16, págs. 99-115.

Johansen, Leif (1977) *Lectures on Macroeconomic Planning*, Ámsterdam: North-Holland.

Johnsen, Jon y Jørg Mørland (1992) «Depot preparations of disulfiram: experimental and clinical results», *Acta Psychiatrica Scandinavica*, 86(S369), noviembre, págs. 27-30.

Korsgaard, Christine M. (1996) *The Sources of Normativity*, Cambridge (Inglaterra) y Nueva York: Cambridge University Press [*Las fuentes de la normatividad*, México: Universidad Nacional Autónoma de México, 2000].

La Bruyère, Jean de (1990) *Les Caractères, ou Les Mœurs de ce siècle*, edición de R. Garapon, París: Bordas, col. «Classiques Garnier» [*Los caracteres o las costumbres de este siglo*, Barcelona: Edhasa, 2004].

La Rochefoucauld (1982) *Maxims*, traducción de Leonard Tancock, Londres: Penguin.

Loewenstein, George y Jon Elster (eds.) (1992) *Choice over Time*, Nueva York: Russell Sage Foundation.

Mackie, Gerry (1996) «Ending footbinding and infibulation: a convention account», *American Sociological Review*, 61(6), diciembre, págs. 999-1017.

Margolis, Howard (1982) *Selfishness, Altruism, and Rationality: A Theory of Social Choice*, Cambridge (Inglaterra) y Nueva York: Cambridge University Press.

Mele, Alfred R. (2001) *Self-Deception Unmasked*, Princeton: Princeton University Press.

Miller, Dale T. (1999) «The norm of self-interest», *American Psychologist*, 54(12), diciembre, págs. 1053-60.

Montaigne, Michel de (1991) *The Essays of Michel de Montaigne*, traducción de M. A. Screech, Londres: Allen Lane.

Neurath, Otto (1913) «Die Verirrten des Cartesius und das Auxiliarmotiv: zur Psychologie des Entschlusses», *Jahrbuch der Philosophischen Gesellschaft an der Universität Wien*, Leipzig: Johann Ambrosius Barth.

Osiatynski, Wiktor (1997) *Alcoholism: Sin or Disease?*, Varsovia: Stefan Batory Foundation.

Parfit, Derek (1984) *Reasons and Persons*, Nueva York y Oxford: Oxford University Press/Clarendon Press

[*Razones y personas*, Madrid: A. Machado Libros, 2004].

Pascal, B. (1995) *Pensées*, traducción de A. J. Krailsheimer, Londres: Penguin.

Proust, Marcel (1988*a*) *À l'ombre des jeunes filles en fleur*, en *À la recherche du temps perdu*, vol. 2, París: Gallimard, col. «Bibliothèque de la Pléiade» [*En busca del tiempo perdido*, vol. 2, *A la sombra de las muchachas en flor*, Buenos Aires: Losada, 2007].

(1988*b*) *La Prisonnière*, en *À la recherche du temps perdu*, vol. 3, París: Gallimard, col. «Bibliothèque de la Pléiade» [*En busca del tiempo perdido*, vol. 5, *La prisionera*, Buenos Aires: Losada, 2005].

Rudenfeld, Jed (2001) *Freedom and Time: A Theory of Constitutional Self-Government*, New Haven: Yale University Press.

Schlicht, Ekkehart (1984) «Die emotive und die kognitive Gerechtigkeitsauffassung», *Ökonomie und Gesellschaft: Wohlfahrt und Gerechtigkeit*, 2, págs. 141-54.

Sen, A. (1973) «Behaviour and the concept of preference», *Economica*, 40, págs. 241-59.

Serck-Hanssen, Jan (1971) «Subsidiering av kapital i utbyggingsområdene», *Statsøkonomisk Tidsskrift*, 84, 140-61.

Stendhal (1980) *De l'amour*, París: Gallimard [*Del amor*, Madrid: Alianza, 1968].

Sugden, Robert (1984) «Reciprocity: the supply of public goods through voluntary contributions», *The Economic Journal*, 94(376), diciembre, págs. 772-87.

Tocqueville, Alexis de (1992) *De la démocratie en Amérique*, en *Œuvres complètes*, vol. 2, París: Gallimard, col. «Bibliothèque de la Pléiade» [*La democracia en América*, México: Fondo de Cultura Económica, 1957].

(2004) *Œuvres complètes*, vol. 3, París: Gallimard, col. «Bibliothèque de la Pléiade».

Veyne, Paul (1976) *Le Pain et le cirque*, París: Seuil.

Watzlawick, Paul (1986) *Le Langage du changement: éléments de communication thérapeutique*, París: Seuil [*El lenguaje del cambio: nueva técnica de la comunicación terapéutica*, Barcelona: Herder, 1992].

Wiener, Jonathan B. (1998) «Managing the iatrogenic risks of risk management», *Risk: Health, Safety and Environment*, 9, págs. 39-82.

Williams, Bernard A. O. (1981) «Internal and external reasons», en *Moral Luck: Philosophical Papers, 1973-1980*, Cambridge (Inglaterra) y Nueva York: Cambridge University Press [«Razones internas y externas», en *La fortuna moral: ensayos filosóficos, 1973-1980*, México: Universidad Nacional Autónoma, 1993, págs. 131-45].

Winter, Sidney G. (1964) «Economic "natural selection" and the theory of the firm», *Yale Economic Essays*, 4, págs. 225-72.

Las lecciones inaugurales del Collège de France

Desde su fundación en 1530, el Collège de France tiene por misión principal la de enseñar, no saberes constituidos, sino «el saber en construcción»: la investigación científica e intelectual misma. Allí, los cursos son gratuitos y están abiertos a todos, sin inscripción ni entrega de diplomas.

De conformidad con su divisa («*Docet omnia*», «Enseña todas las cosas»), el Collège de France está organizado en cincuenta y dos cátedras que abarcan una vasta serie de disciplinas. Anualmente se dictan, además, una cátedra europea, una cátedra internacional, una cátedra de creación artística y, a partir de 2006, una cátedra de innovación tecnológica.

Los profesores son elegidos libremente por sus pares, en función de la evolución de las ciencias y de los conocimientos. A la llegada de cada nuevo profesor se crea una cátedra, que puede retomar al menos en parte la herencia de una anterior, o bien instaurar una nueva enseñanza.

La primera clase de un nuevo profesor es su lección inaugural.

Solemnemente pronunciada en presencia de sus colegas y de considerable público, esa lección le brinda al disertante la oportunidad de situar sus trabajos y su enseñanza con respecto a los de sus predecesores y a los desarrollos más recientes de la investigación.

Las lecciones inaugurales no sólo trazan un cuadro del estado de nuestros conocimientos y contribuyen así a la historia de cada disciplina, sino que nos introducen también en el taller del estudioso y del investigador. Muchas de ellas constituyeron, en su ámbito y en su tiempo, acontecimientos destacados e incluso resonantes.

Estas lecciones están dirigidas a un vasto público ilustrado, deseoso de comprender mejor las evoluciones de la ciencia y de la vida intelectual contemporáneas.

FILOSOFÍA

Razón y racionalidad
Jon Elster

Jon Elster, uno de los filósofos políticos más importantes del mundo, se destaca por sus reflexiones sobre la razón y la racionalidad y el papel de ambas en la política y en la vida pública. Este breve libro significa, a la vez, una cristalización y un avance de su obra, al tender un puente entre los filósofos, que para evaluar el comportamiento humano desde un punto de vista normativo utilizan la idea de razón, y los especialistas en ciencias sociales, que para explicar aquel se valen de la idea de racionalidad. En lugar de esos abordajes, Elster propone para el estudio del comportamiento un marco conceptual unificado.

Sobre la base tanto de moralistas clásicos como del saber moderno, y a través de abundantes ilustraciones históricas y contemporáneas, Razón y racionalidad *marca un nuevo desarrollo en el pensamiento de Elster y nos proporciona, al mismo tiempo, una breve, elegante y accesible introducción a su obra.*

Jon Elster ocupa la Cátedra de Racionalidad y Ciencias Sociales en el Collège de France. Entre sus libros recientes se hallan La explicación del comportamiento social *y* Rendición de cuentas: la justicia transicional en perspectiva histórica.

ISBN 978-950-518-360-9

Cubierta: Diseño A

Amorrortu/*editores*